

# Coherence and measurement in quantum thermodynamics

Philipp Kammerlander\*

*Institute for Theoretical Physics, ETH Zurich, Wolfgang-Pauli-Strasse 27, 8093 Zurich, Switzerland.*

Janet Anders†

*Department of Physics and Astronomy, University of Exeter, Stocker Road, EX4 4QL, United Kingdom.*

Thermodynamics is a highly successful macroscopic theory widely used across the natural sciences and for the construction of everyday devices, from car engines and fridges to power plants and solar cells. With thermodynamics predating quantum theory, research now aims to uncover the thermodynamic laws that govern finite size systems which may in addition host quantum effects. Here we identify information processing tasks, the so-called “projections”, that can only be formulated within the framework of quantum mechanics. We show that the physical realisation of such projections can come with a non-trivial thermodynamic work *only* for quantum states with coherences. This contrasts with information erasure, first investigated by Landauer, for which a thermodynamic work cost applies for classical and quantum erasure alike. Implications are far-reaching, adding a thermodynamic dimension to measurements performed in quantum thermodynamics experiments, and providing key input for the construction of a future quantum thermodynamic framework. Repercussions are discussed for quantum work fluctuation relations and thermodynamic single-shot approaches.

## I. INTRODUCTION

When Landauer argued in 1961 that any physical realisation of erasure of information has a fundamental thermodynamic work cost he irrevocably linked thermodynamics and information theory [1]. A practical consequence of this insight is that all computers must dissipate a minimal amount of heat in each irreversible computing step, a threshold that is becoming a concern with future computer chips entering atomic scales. The treatment of general *quantum* information processing tasks within the wider framework of quantum thermodynamics has only recently begun. Theoretical breakthroughs include the characterisation of the efficiency of quantum thermal engines [2–4] and the extension of widely used classical non-equilibrium fluctuation theorems to the quantum regime [5, 6]. A new thermodynamic resource theory [7] has led to the discovery of a *set* of second laws that replaces the standard macroscopic second law for finite size systems [8, 9]. These results have substantially advanced our understanding of nanoscale thermodynamics, however putting a finger on what is genuinely “quantum” in quantum thermodynamics has remained a challenge. Quantum mechanics differs from classical mechanics in at least three central aspects: the special nature of measurement, the possibility of a quantum system to be in a superposition and the existence of quantum correlations. The thermodynamic energy needed to perform a (selective) measurement has been investigated [10] and the total work for a closed thermodynamic measurement cycle explored [11]. The catalytic role of quantum superposition states when used in thermal operations has been uncovered [12] and it has been shown that work can

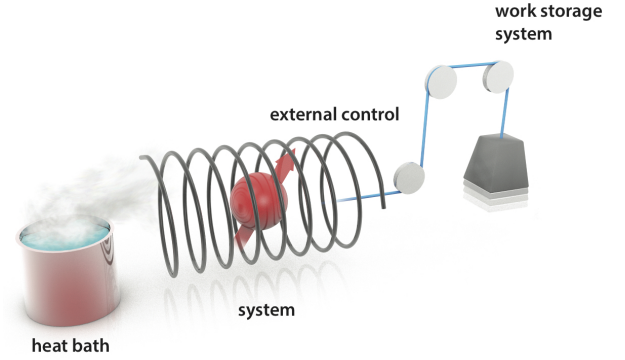


FIG. 1: **Thermodynamic setting.** A system, depicted as a spin, interacts with a heat bath at temperature  $T$ , with which it exchanges *heat*, and with controlled energy sources, illustrated as coil and weight, with which it exchanges *work*. Work drawn from the system can be collected in a work storage system (weight) for future use.

be drawn from quantum correlations [13, 14] in a thermodynamic setting, see Fig. 1. In particular, del Rio *et al.* [14] showed that contrary to Landauer’s principle, it is possible to *extract* work while performing erasure of a system’s state when the system is correlated to a memory. This can occur if and only if the initial correlations imply a negative conditional entropy, a uniquely quantum feature. The thermodynamic process does however now require operation on degrees of freedom external to the system, i.e. the memory’s.

Our motivation is here to shed light on the implications of performing a measurement on a quantum state that has coherences. We will consider this task in the thermodynamic setting of Landauer’s erasure, involving a heat bath at fixed temperature  $T$  and operation on  $N \rightarrow \infty$  uncorrelated and identically prepared copies of the system (i.i.d. limit). This is of interest in the

\*Electronic address: kammerlander@phys.ethz.ch

†Electronic address: janet@qipc.org

context of the quantum Jarzynski equality, for example, and will also be central for experiments testing quantum thermodynamic predictions in the future. To tackle this question we define the information-theoretic “projection”  $\rho \rightarrow \eta^{\mathcal{P}} := \sum_k \Pi_k^{\mathcal{P}} \rho \Pi_k^{\mathcal{P}}$  for a given initial quantum state  $\rho$  and a complete set of mutually orthogonal projectors  $\{\Pi_k^{\mathcal{P}}\}_k$ . Such state transformation can be seen as analogous to the state transfer of erasure,  $\rho \rightarrow |0\rangle$ , to a blank state  $|0\rangle$ . Physically, this projection can be interpreted as the result of an unread, or unselective [15], measurement of an observable  $\mathcal{P}$  that has eigenvector projectors  $\{\Pi_k^{\mathcal{P}}\}_k$ . In an unselective measurement the individual measurement outcomes are not recorded and only the statistics of outcomes is known. In the literature the implementation of unselective measurements is often not specified, although it is typically thought of as measuring individual outcomes, e.g. with a Stern-Gerlach experiment, see Fig. 2a, followed by mixing. The crux is that the information-theoretic projection  $\rho \rightarrow \eta^{\mathcal{P}}$  can be implemented in many physical ways. The associated thermodynamic heat and work will differ depending on *how* the projection was done and we will refer to the various realisations as “thermodynamic projection processes”. One possibility is decohering [16] the state in the so-called pointer basis,  $\{\Pi_k^{\text{pointer}}\}_k$ , a thermodynamic process where an environment removes coherences in an uncontrolled manner resulting in no associated work. In general it is possible to implement the state transfer in a finely controlled fashion achieving optimal thermodynamic heat and work values.

## II. MAIN RESULT

Of particular importance in thermodynamics is the projection  $\rho \rightarrow \eta^H$  of the system’s initial state  $\rho$  onto the set of energy eigenstates  $\{\Pi_k^H\}_k$  of the system’s Hamiltonian  $H = \sum_k E_k \Pi_k^H$  with  $E_k$  the energy eigenvalues. Here the state’s off-diagonals with respect to the energy eigenbasis are removed - a state transformation that is frequently employed in quantum thermodynamic derivations and referred to as “dephasing” or “measuring the energy”. Our key observation is that there exists a thermodynamic projection process realising this transformation and allowing to draw from the quantum system a non-trivial *optimal average work* of

$$\langle W_{\text{opt}} \rangle = k_B T (S(\eta^H) - S(\rho)). \quad (1)$$

Here  $T$  is the temperature of the heat bath with which the system is allowed to interact, see illustration Fig. 1,  $k_B$  is the Boltzmann constant and  $S$  is the von Neumann entropy. Crucially, this work is strictly positive for quantum states with coherences. Extending the key observation to general projections  $\rho \rightarrow \eta^{\mathcal{P}}$  one finds that optimal thermodynamic projection processes can be implemented that allow to draw an average work of

$$\langle W_{\text{opt}} \rangle = k_B T (S(\eta^{\mathcal{P}}) - S(\rho)) - \text{tr} [H (\eta^{\mathcal{P}} - \rho)], \quad (2)$$

where an additional internal energy change term appears.

The optimal work values stated in Eqs. (1) and (2) are valid for processes applied to classical and quantum states alike. While for a classical ensemble the entropy change,  $\Delta S^{\mathcal{P}} = S(\eta^{\mathcal{P}}) - S(\rho)$ , will be zero this is not so in the general quantum situation, where initial non-diagonal quantum states result in a strictly positive entropy change [17]. We note that while the optimal work values are in principle attainable, practical implementations may be suboptimal resulting in a reduced work gain or a higher work cost. The physical meaning of  $\Delta S^{\mathcal{P}}$  can be grasped by considering a lower bound [18] on it,  $\Delta S^{\mathcal{P}} \geq \frac{1}{2} \|\rho - \frac{1}{d} \mathbb{I}\|_2^2 \Delta A^{\mathcal{P}}$ , see Appendix E. Here  $d$  is the dimension of the system and  $\|\cdot\|_2$  denotes the Hilbert-Schmidt norm. The first factor quantifies the distance of the initial state from the fully mixed state, while the second factor,  $\Delta A^{\mathcal{P}}$ , quantifies the angle between the diagonal basis of  $\rho$  and the projection basis  $\{\Pi_k^{\mathcal{P}}\}_k$ . These terms correspond to incoherent and coherent mixing contributions. The entropy change is non-trivially bounded only if the initial state is not an incoherent mixture with respect to that basis. The entropy bound is the largest for pure initial states whose basis is mutually unbiased with respect to  $\{\Pi_k^{\mathcal{P}}\}_k$ . In this case the optimal entropy change is  $\Delta S^{\mathcal{P}} = k_B T \ln d$ .

One may wonder where the work has gone to. There are two equivalent approaches to the accounting of work. In the present analysis the focus is on the work that the system exchanges, as done in statistical physics [6, 19–22]. In this approach it is often not explicitly mentioned where the work goes to, but the only place work can go to are the externally controlled energy sources. Similarly, the heat, i.e. the energy change minus the work, is established implicitly. For example, in the experimental realisation of classical Landauer erasure with a colloidal silica bead trapped in an optical tweezer [21], the dissipated heat of erasure was calculated by knowing the applied tilting forces and integrating over the bead’s dynamics. The second approach is to collect work in a separate work storage system [23], as illustrated by the weight in Fig. 1 and detailed in Appendix C. Both the implicit and the explicit treatment of work are equivalent in the sense that the results obtained in one approach can be translated into the other.

The thermodynamic assumptions made to prove Eq. (2) are congruent with current literature [1, 23–25]; specifically they are: (T0) an isolated system is a system that only exchanges work and not heat; (T1) the validity of the *first law* relating the internal energy change,  $\Delta U$ , of the system during a process to its average heat absorbed and work drawn,  $\Delta U = \langle Q^{\text{abs}} \rangle - \langle W \rangle$ ; (T2) the validity of the *second law* relating the system’s entropy change to its average absorbed heat,  $k_B T \Delta S \geq \langle Q^{\text{abs}} \rangle$ , when interacting with a bath at temperature  $T$ , with equality attainable by an optimal process; (T3) the thermodynamic entropy to be equal to the von Neumann entropy in equilibrium as well as out-of-equilibrium,  $S_{\text{th}} = S_{\text{vN}}$ . In addition we make the following standard quan-

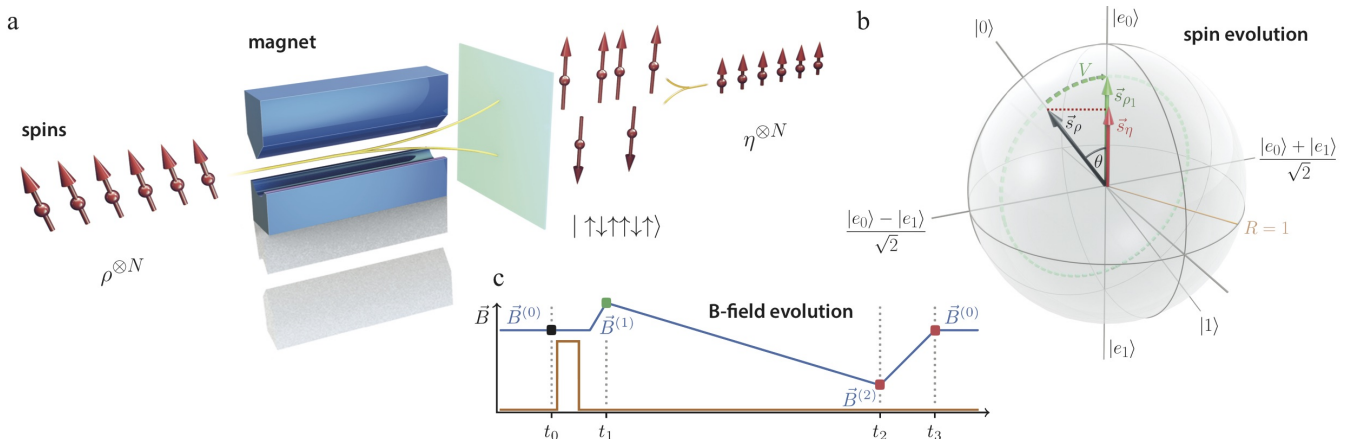


FIG. 2: **Two physical realisations of a projection process.** **a**,  $N$  identically prepared spin 1/2 particles in state  $\rho^{\otimes N}$  pass a Stern-Gerlach magnet and a screen after which they emerge in either the spin-up or the spin-down beam. Recombining the two beams mixes the spins to the final state  $\eta^{\otimes N}$  for  $N \rightarrow \infty$ . Illustration of the spin example, showing the state evolution in **b** and the B-field evolution in **c**. The poles in the Blochsphere (**b**) are the energy eigenstates  $|e_0\rangle$  and  $|e_1\rangle$  that are aligned and anti-aligned with an externally applied B-field (indicated in blue in **c**), which initially is  $\vec{B}^{(0)}$  (black point in **c**). In the first step the Blochvector  $\vec{s}_\rho$  (black arrow in **b**) of Emmy's initial state  $\rho$  is rotated on the green-dashed circle to  $\vec{s}_{\rho_1}$  (green arrow in **b**). The unitary rotation  $V$  required for this step can be realised by applying a microwave pulse creating an additional B-field (indicated in orange in **c**) in the direction orthogonal to the plane of the green circle. At the end of the first step the pulse is turned off and the external B-field is adjusted to  $\vec{B}^{(1)}$  (green point in **c**). The second step shortens  $\vec{s}_{\rho_1}$  to  $\vec{s}_\eta$  (red arrow in **b**), the Blochvector of  $\eta$  (superscripts  $H$  have been omitted). The external B-field (blue in **c**) decreases slowly to  $\vec{B}^{(2)}$  (red point at  $t_2$  in **c**). In the last step the B-field quickly returns to its initial value,  $\vec{B}^{(0)}$  (red point at  $t_3$  in **c**), while the state remains  $\eta$ . The angle between the Blochvectors of  $\rho$  and  $\eta$  is indicated by  $\theta$ .

tum mechanics assumptions: (Q0) an isolated system evolves unitarily; (Q1) control of a quantum system includes its coherences. Details of the proof are in Appendix A. We note that in the single-shot setting whole families of second laws apply [8, 9] that differ from (T2) stated above. However, in the limit of infinitely many independent and identically prepared copies of the system these collapse to the standard second law, (T2), on the basis of which Eq. (2) is derived.

From the information-theory point of view the projections considered here constitute just one example of the larger class of trace-preserving completely positive (TPCP) maps characterising quantum dynamics. Of course, all TPCP maps can be interpreted thermodynamically with the assumptions stated above, resulting in an optimal average work given by a free energy difference. Erasure is another such map whose study forged the link between information theory and thermodynamics. The benefit of discussing “projections” here lies in the insight that this focus provides: it uncovers that coherences offer the potential to draw work making it a genuine and testable quantum thermodynamic feature. This work is non-trivial even when the thermodynamic process is operated on the system alone, not involving any side-information [14] stored in other degrees of freedom.

### III. EXAMPLE

To gain a detailed understanding of thermodynamic projection processes that give the optimal work stated in Eq. (1) we now detail one such process for the example of a spin-1/2 particle (qubit), see illustration in Fig. 2b and 2c as well as Appendix B. This process consists of a unitary evolution, a quasi-static evolution and a quench [25], and it is optimal for any finite-dimensional quantum system as shown in Appendix D. An experimentalist, Emmy, prepares the spin in a state  $\rho = a|0\rangle\langle 0| + (1-a)|1\rangle\langle 1|$  ( $a \geq \frac{1}{2}$  w.l.o.g.) exposed to an external magnetic field  $\vec{B}^{(0)}$  which she controls. The Hamiltonian associated with the system is  $H = -E|e_0\rangle\langle e_0| + E|e_1\rangle\langle e_1|$  where the energy difference between the aligned ground state,  $|e_0\rangle$ , and anti-aligned excited state,  $|e_1\rangle$ , is given by  $2E = |\vec{\mu}||\vec{B}^{(0)}|$  with  $\vec{\mu}$  the spin's magnetic moment. Importantly, in general the spin state's basis,  $\{|0\rangle, |1\rangle\}$ , are superpositions with respect to the energy eigenbasis,  $|0\rangle = \alpha^*|e_0\rangle + \beta^*|e_1\rangle$  and  $|1\rangle = \beta|e_0\rangle - \alpha|e_1\rangle$  with  $|\alpha|^2 + |\beta|^2 = 1$ . For the optimal implementation of the projection  $\rho \rightarrow \eta^H = \sum_{k=0,1} |e_k\rangle\langle e_k| \rho |e_k\rangle\langle e_k|$  Emmy now proceeds with the following three steps.

Firstly, she isolates the spin from the bath and modifies external magnetic fields to induce a unitary rotation,  $V = |e_0\rangle\langle 0| + |e_1\rangle\langle 1|$ , of the spin into the energy basis. In nuclear magnetic resonance (NMR) [26] and pulsed electron spin resonance (ESR) experiments [27] such rotations are routinely realised by radio-frequency

and microwave pulses respectively, as evidenced by Rabi oscillations. The power, duration and phase of such a pulse would be chosen to generate the spin-rotation along the green circle until the desired unitary  $V$  is achieved. In the same step Emmy adjusts the strength of the external B-field such that the spin state  $\rho_1 = V \rho V^\dagger$  is Boltzmann-distributed at temperature  $T$  with respect to the energy gap of the Hamiltonian at the end of the step,  $H^{(1)}$ . In NMR or ESR the B-field magnitude is tuned quickly on the  $T_1$  timescale to achieve the desired energy gap. In the second step, Emmy wants to implement a quasi-static evolution of the spin that is now thermal. She brings the spin in contact with the heat bath at temperature  $T$  and quasi-statically adjusts the magnitude of the external B-field allowing the spin state to thermalise at all times. The final B-field,  $\vec{B}^{(2)}$ , is chosen such that the final thermal state becomes  $\eta^H$ . In ESR this step can be realised by changing the external B-field slowly on the  $T_1$  timescale so that the spin continuously equilibrates with its environment. Finally, Emmy isolates the spin from the environment and quickly changes the B-field to its original magnitude while the state remains  $\eta^H$ .

During Step 1 and 3 the system was isolated and the average work drawn is thus just the average energy change. During Step 2 the average work is the equilibrium free energy difference between the final and initial thermal states at temperature  $T$ , see Appendix B for details. In NMR/ESR the work contributions drawn from the spin system are done on the external B-field and the microwave mode. This could be detected by measuring the stimulated emission of photons in the microwave mode or observing current changes induced by the spins dynamics [26, 27]. The overall thermodynamic process has now brought the spin from a quantum state with coherences,  $\rho$ , into a state without coherences,  $\eta^H$ , while keeping the average energy of the spin constant. The net work drawn during the three steps adds up to  $\langle W \rangle = k_B T (S(\eta^H) - S(\rho))$  showing the attainability of the optimum stated in Eq. (1) for the spin-1/2 example. We note that Eq. (1) is also the maximal work that can be extracted from a *qubit* state  $\rho$  under *any* transformation of the system that conserves its average energy,  $U := \text{tr}[H \rho]$ , i.e. for qubits  $\eta^H$  is the optimal final state under this condition.

We emphasise that this optimal implementation involves a finely tuned and controlled operation that relies on knowledge of the initial state  $\rho$ . This is akin to the situation considered in [14] where knowledge of the initial global state of system and memory is required for optimal erasure with side-information. It is important to distinguish this situation from that of Maxwell demon's who has access to knowledge of the individual micro-states  $\Pi_k^P$  that make up the ensemble state  $\eta^P$ , and who uses it to beat the second law [28]. In the scenario considered here there is no knowledge of the individual micro-states  $\Pi_k^P$  and the process does not violate the second law, on the contrary, it is derived from it.

## IV. IMPLICATIONS

### A. Single-shot analysis

The preceding discussion concerned the *average* work that can be drawn when operating on an ensemble of  $N \rightarrow \infty$  independent spins. This scenario contrasts with the single shot situation considered in a number of recent publications [8, 14, 29, 30]. In particular, two major frameworks [29, 30] have recently been put forward to identify optimal *single-shot* work extraction and work cost of formation in the quantum setting. These frameworks rely on a resource theory approach [7] and make use of min- and max-relative entropies that originate from one-shot information theory. The optimal work extraction schemes of these frameworks require non-diagonal states to be decohered first to become diagonal in the energy basis. This decoherence step is assumed to not have an associated single-shot work. However, the present analysis of energy basis projections showed that thermodynamic projection processes can yield positive average work, see Eq. (1). Therefore one may expect a positive work for removing coherences from a state  $\rho$  in the single-shot setting, too. Since our focus is the  $N \rightarrow \infty$  limit we will not aim to construct the single-shot case. Nevertheless, to establish a notion of consistency between single-shot results [29, 30] and the average analysis presented here we now separate the projection into a diagonal part that can be analysed in the single-shot framework and a non-diagonal part that can be analysed in the average framework. One possible decomposition of  $\rho \rightarrow \eta^H$  is the split in three steps each starting and ending with Hamiltonian  $H$ :  $\rho \xrightarrow{a} \rho_1 \xrightarrow{b} \tau^H \xrightarrow{c} \eta^H$ . Here  $\rho_1$  is the rotated state defined above and  $\tau^H = e^{-H/k_B T} / \text{tr}[e^{-H/k_B T}]$  is the thermal state for the Hamiltonian  $H$  at temperature  $T$ . We can now use a single-shot analysis [30] for Steps  $b$  and  $c$  that involve only states diagonal in the energy basis, giving a single-shot work contribution of  $k_B T \ln 2 (D_{\min}(\rho_1 || \tau^H) - D_{\max}(\eta^H || \tau^H))$ , see Appendix F. Here  $D_{\min}$  and  $D_{\max}$  are the min- and max-relative quantum entropies, respectively. Taking the limit of  $N \rightarrow \infty$  copies for Steps  $b$  and  $c$  and adding the average work contribution for the initial non-diagonal rotation  $a$ ,  $\langle W^{(a)} \rangle = -\text{tr}[(\rho_1 - \rho) H]$ , one indeed recovers the optimal average work as stated in Eq. (1). After making public our results very recently a paper appeared [31] that derives the work that can be extracted when removing coherences in a single-shot setting. These results are in agreement with Eq. (1) and reinforce the above conclusion that coherences are a fundamental feature distinguishing quantum from classical thermodynamics.

### B. Quantum fluctuation relations

The key observation was that thermodynamic projection processes can have a non-trivial work and heat. An-

other instance where this has interesting repercussions is the quantum Jarzynski equality [5, 6]. This is a generalisation of the prominent classical fluctuation relation valid for general non-equilibrium processes, which has been used to measure the equilibrium free energy surface inside bio-molecules by performing non-equilibrium pulling experiments [19]. The quantum version has recently been tested for the first time in a nuclear magnetic resonance experiment [26]. The quantum Jarzynski relation,  $\langle e^{W/(k_B T)} \rangle = e^{-\Delta F/(k_B T)}$ , links the fluctuating work,  $W$ , drawn from a system in individual runs of the same non-equilibrium process, with the free energy difference,  $\Delta F$ , of the thermal states of the final and initial Hamiltonian, see Appendix G. In its derivation a system initially in a thermal state  $\rho_0$  with respect to Hamiltonian  $H^{(0)}$  at temperature  $T$  is first measured in the energy basis of  $H^{(0)}$ . The Hamiltonian is then varied in time ending in  $H^{(\tau)}$  generating a unitary evolution,  $V$ , of the system, see Fig. 3a. A second measurement, in the energy basis of  $H^{(\tau)}$ , is then performed to establish the final fluctuating energy. For each run the difference of the two measured energies has been associated with the fluctuating work [6],  $\Delta E = -W$ . The experiment is repeated, each time producing a fluctuating work value. On average the work extracted from the system during the quantum non-equilibrium process turns out to be  $\langle W^{\text{unitary}} \rangle = U(\rho_0) - U(\rho_\tau)$  where  $\rho_\tau := V \rho_0 V^\dagger$  is the ensemble's state after the unitary evolution, and similarly the average exponentiated work is calculated. The above identification  $W := -\Delta E$  was made assuming that the system undergoes a unitary process with no heat dissipation. However, the need to acquire knowledge of the system's final energies requires the second measurement. The ensemble state is thus further altered from  $\rho_\tau$  to  $\eta_\tau$ , the state  $\rho_\tau$  with any coherences in the energy basis of  $H^{(\tau)}$  removed. This step is not unitary - during the projection  $\rho_\tau \rightarrow \eta_\tau$  the system may absorb heat,  $\langle Q^{\text{abs}} \rangle$ , indicated in Fig. 3b, whose value depends on *how* the process is conducted. Thus, while the energy difference for the projection is zero,  $U(\eta_\tau) - U(\rho_\tau) = 0$ , for states  $\rho_\tau$  with coherences the entropy difference is not trivial,  $S(\eta_\tau) - S(\rho_\tau) = \langle Q_{\text{opt}}^{\text{abs}} \rangle / (k_B T) \geq 0$ . This implies that in an experimental implementation of the Jarzynski relation the work done by the system on average can be more than previously thought,  $\langle W_{\text{opt}} \rangle = \langle W^{\text{unitary}} \rangle + k_B T (S(\eta_\tau) - S(\rho_\tau))$ . We conclude that the suitability of identifying  $W = -\Delta E$ , and hence the validity of the quantum Jarzynski *work* relation, depends on the details of the physical process that implements the second measurement. This conclusion is not at odds with previous experiments [26] which showed nature's agreement with  $\langle e^{-\Delta E/(k_B T)} \rangle = e^{-\Delta F/(k_B T)}$ , involving the average of the exponentiated measured fluctuating energy.

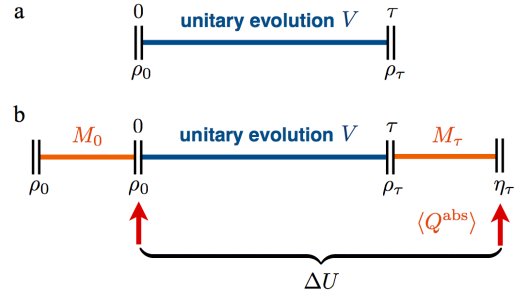


FIG. 3: **Dynamical steps in a quantum fluctuation experiment.** **a**, The quantum Jarzynski relation is described as characterising the non-equilibrium work of processes that start in a thermal state  $\rho_0$  and evolve unitarily ( $V$ ), driven by a changing Hamiltonian, reaching the final state  $\rho_\tau$  at time  $\tau$ . This unitary process has no heat contribution. **b**, Illustration of three steps that are assumed in mathematical derivations of the quantum Jarzynski relation [5, 6]: initial energy measurement of  $H^{(0)}$  indicated by  $M_0$ , unitary evolution, and final energy measurement of  $H^{(\tau)}$  indicated by  $M_\tau$ . The ensemble state evolves here from  $\rho_0$  to  $\rho_\tau$  and then to  $\eta_\tau$ , the state  $\rho_\tau$  with its coherences removed. The observed average energy difference  $\Delta U = U(\eta_\tau) - U(\rho_0)$  encompasses both, the unitary process and the second projection process, and can in general contain a heat contribution  $\langle Q^{\text{abs}} \rangle$ , in contrast to **a**.

### C. Correlated systems

It is insightful to extend the thermodynamic analysis of projections to correlated systems. An experimenter may have access not only to the system  $S$  but also the auxiliary systems  $A$  with which  $S$  is correlated [14]. She can then perform a global operation,  $\rho^{SA} \rightarrow \eta^{SA}$ , that implements a projection  $\{\Pi_k^P\}_k$  locally on the system  $S$ , i.e.  $\rho^S := \text{tr}_A[\rho^{SA}] \rightarrow \eta^S := \text{tr}_A[\eta^{SA}] = \sum_k \Pi_k^P \rho^S \Pi_k^P$ , while leaving the reduced state of the auxiliary system unchanged, i.e.  $\rho^A := \text{tr}_S[\rho^{SA}] = \text{tr}_S[\eta^{SA}]$ . By doing so the experimenter can optimally draw the overall work  $\langle W_{\text{opt}} \rangle = k_B T \Delta S^P - \Delta U^P$ , where  $\Delta S^P$  is the entropy change for the state of system+auxiliary and  $\Delta U^P$  is still the energy change of the system *alone*. This quantity can be re-written as the sum of two terms:  $\langle W_{\text{opt}} \rangle$ , the extractable work when operating on the system alone given in Eq. (2), and  $\delta^P(A : S)$ , a positive term quantifying the quantum correlations between  $S$  and  $A$ , see Appendix H. The latter contribution was previously identified in an inspiring paper by Zurek [13]. It depends on the choice of projectors and is related to, but broader than, quantum discord [32] which is optimised over all possible projectors. This means that even states of system and auxiliary that can be considered classically correlated (i.e. no discord) provide an advantage for drawing work contrasting with the erasure process where this only occurs for highly entangled states [14]. The gap between these two sets of correlated states is an intriguing fact and calls for further exploration of the link between thermodynamics and information theory in the quantum regime.



## V. CONCLUSIONS

To conclude, erasure is not the only irreversible information processing task – in the quantum regime a second fundamental process exists that mirrors Landauer’s erasure. In contrast to the minimum heat limit of erasure, thermodynamic projection processes have a maximum work limit. While the former is non-zero for the erasure of classical *and* quantum bits, optimal thermodynamic projection processes have a non-zero work *only* when applied to quantum states with coherences. The optimal average work stated in Eqs. (1) and (2) constitutes an experimentally accessible quantum thermodynamic prediction. Future experiments testing this optimal work may be pursued with current setups, for instance with NMR/ESR techniques [26, 27] or single atoms [33, 34], and promise to be accessible with other platforms entering the quantum regime, such as single electron boxes [22]. Experiments will be limited by practical constraints, such as achieving a quasistatic process and obtaining the maximum work for pure states which may require, for instance, very large B-fields.

The derivation of the optimal work value is mathematically straightforward, just like that of Landauer’s principle. The result’s significance is that it opens new avenues of thought and provides key input for the construction of a future quantum thermodynamic framework. For example, the developed approach opens the door to investigate the connection between microscopic statistical physics and macroscopic thermodynamics in the quantum regime. While it is straightforward to identify the thermodynamic work of quantum processes involving macroscopic ensembles, what is needed is a microscopic concept of work that when averaged, gives the correct macroscopic work. The microscopic work concept should be valid for general (open) quantum processes and quantum states (including coherences), and only require access to properties of the system. While single-shot approaches have discarded coherences [29, 30], fluctuating work approaches cannot be applied directly to a system undergoing open quantum evolution [20].

The observation is also important from the experimental perspective as testing quantum thermodynamic predictions will involve measurement – a projection process. We have argued that measurements, such as those required in establishing the Jarzynski equality, are not necessarily thermodynamically neutral. Indeed, they can be implemented in different physical ways and in general play an active role in thermodynamics, contributing a non-zero average heat and work. This new perspective gives physical meaning to the change of entropy in the debated quantum measurement process - it provides a capacity to draw work. Specifically, work can be drawn when *coherences* of a state are removed during an unselective measurement.

Finally, it is apparent that optimal thermodynamic projection processes require use of knowledge of the initial state  $\rho$ , i.e. its basis and eigenvalues. One may

be inclined to exclude use of such knowledge, particularly when considering projections in the context of measurement which is often associated with the acquisition of knowledge. Such restriction would necessarily affect the set of assumptions (T0-T3, Q0-Q1) in the quantum regime. These could be changed, for example, to the second law not being possible to saturate (cf. T2) or to choosing a new quantum non-equilibrium entropy that only considers the state’s diagonal entries (cf. T3). The latter would mean a departure from standard quantum information theory where entropies are basis-independent. Thus whichever approach one takes - not making or making a restriction - quantum coherences will contribute a new dimension to thermodynamics. They either lead to non-classical work extraction or they alter the link between information theory and thermodynamics in the quantum regime. The line drawn here between the assumptions (T0-T3, Q0-Q1) and results (Eqs. (1) and (2)) establishes a frame for this possibility to be investigated.

## Acknowledgments

We thank T. Deesuan, M. Wolf and R. Renner, G. Morley, R. Uzdin and D. Reeb for insightful discussions and J. Gemmer, R. Renner, S. Horsley and T. Philbin for critical reading of the manuscript. P.K. acknowledges support from the Swiss National Science Foundation (through the National Centre of Competence in Research ‘Quantum Science and Technology’) and the European Research Council (grant 258932). J.A. is supported by the Royal Society and EPSRC. J.A. thanks the Isaac Newton Institute in Cambridge where part of this work was conceived for the stimulating environment and kind hospitality. This work was supported by the European COST network MP1209.

## Appendix A: Proof of Eq. (2)

Using the first law (T1) the average work drawn in a thermodynamic projection process  $\rho \rightarrow \eta^{\mathcal{P}}$  is simply  $\langle W \rangle = \langle Q^{\text{abs}} \rangle - \Delta U^{\mathcal{P}}$ , where  $\Delta U^{\mathcal{P}}$  is the average energy change for that process. Relating the average heat absorbed by the system during the process to its entropy change one then obtains  $\langle W \rangle \leq k_B T \Delta S^{\mathcal{P}}$  (T2). Here  $\Delta S^{\mathcal{P}}$  is the difference of von Neumann entropies of the system’s state before and after the projection (T3). The average work drawn is thus  $\langle W \rangle \leq k_B T \Delta S^{\mathcal{P}} - \Delta U^{\mathcal{P}}$ , where the entropy change is non-negative and the energy change can be either positive or negative. The stated *optimal* work,  $\langle W_{\text{opt}} \rangle$ , is achieved when the inequality is saturated by an optimal process (T2) the implementation of which may require knowledge of the initial state and control of coherences (Q1). In the special case of a projection onto the energy eigenbasis  $\{\Pi_k^H\}_k$  the internal energy change is zero,  $\Delta U^H = 0$ , and one obtains

Eq. (1).

### Appendix B: Spin example

We here detail the three steps of the optimal thermodynamic projection process of the spin system discussed in the main text. Emmy starts with a spin in state

$$\rho = a |0\rangle\langle 0| + (1-a) |1\rangle\langle 1|, \quad (\text{B1})$$

with Blochvector

$$\vec{s}_\rho := \text{tr}[\rho \vec{\sigma}] = (2a-1) \hat{0}, \quad (\text{B2})$$

where  $\vec{\sigma}$  is the vector of the three Pauli matrices,  $\sigma_1, \sigma_2$  and  $\sigma_3$  and  $\hat{0} = \text{tr}[|0\rangle\langle 0| \vec{\sigma}]$  is the unit vector in the Blochsphere pointing from the origin to the state  $|0\rangle$ , see Fig. 2b. We assume without loss of generality that  $a \geq \frac{1}{2}$ . If this was not the case, the labels  $|0\rangle\langle 0|$  and  $|1\rangle\langle 1|$  should be interchanged. The spin's initial Hamiltonian is given by  $H = -E (\Pi_0^H - \Pi_1^H)$ , where  $\Pi_k^H = |e_k\rangle\langle e_k|$  with  $k = 0, 1$  are the rank-1 projectors onto the two energy eigenstates and  $E > 0$ . This Hamiltonian arises when the spin is exposed to an external magnetic field  $\vec{B}^{(0)}$ . The energy separation of the aligned ground state,  $|e_0\rangle$ , and anti-aligned excited state,  $|e_1\rangle$ , is  $2E = 2|\vec{\mu}| |\vec{B}^{(0)}|$ , where  $\vec{\mu}$  is the magnetic moment of the spin. A general initial state  $\rho$  is not diagonal in the basis  $\{|e_0\rangle, |e_1\rangle\}$ , in other words the spin's eigenstates are superpositions with respect to the energy eigenbasis,  $|0\rangle = \alpha^* |e_0\rangle + \beta^* |e_1\rangle$  and  $|1\rangle = \beta |e_0\rangle - \alpha |e_1\rangle$  with  $|\alpha|^2 + |\beta|^2 = 1$ . The spin's Blochvector,  $\hat{0}$ , is then *not* parallel to the B-field,  $\vec{B}^{(0)}$ . Emmy wants to obtain the state where the coherences with respect to the energy basis  $\{|e_0\rangle, |e_1\rangle\}$  have been removed,

$$\begin{aligned} \eta^H &= \sum_{k=0,1} \Pi_k^H \rho \Pi_k^H \\ &= \sum_{k=0,1} \text{tr}[\Pi_k^H \rho] \Pi_k^H = p \Pi_0^H + (1-p) \Pi_1^H, \end{aligned} \quad (\text{B3})$$

where  $p = \text{tr}[\Pi_0^H \rho]$  is the probability for obtaining outcome  $|e_0\rangle$  in a measurement of  $H$ . Here  $\eta^H$  has the same average energy as the initial state  $\rho$ ,

$$\begin{aligned} \text{tr}[\eta^H H] &= \sum_{k=0,1} \text{tr}[\Pi_k^H \rho \Pi_k^H H] \\ &= \sum_{k=0,1} \text{tr}[\rho \Pi_k^H H \Pi_k^H] = \text{tr}[\rho H], \end{aligned} \quad (\text{B4})$$

where we used orthonormality and completeness of the projectors  $\{\Pi_0^H, \Pi_1^H\}$ . The Blochvector of the final state is defined as

$$\vec{s}_\eta := \text{tr}[\eta^H \vec{\sigma}] = (2p-1) \hat{e}_0 \quad (\text{B5})$$

where  $\hat{e}_0 = \text{tr}[|e_0\rangle\langle e_0| \vec{\sigma}]$  is the unit vector in the Blochsphere pointing from the origin to the state  $|e_0\rangle$ .

Since geometrically the mapping  $\rho \rightarrow \eta^H$  is a projection of  $\vec{s}_\rho$  onto the vertical axis in the Blochsphere, the length of the final Blochvector,  $\vec{s}_\eta$ , is shorter than the initial Blochvector,  $\vec{s}_\rho$ . This shortening is associated with an entropy increase [17]. When describing the process in the following we assume that  $p \geq \frac{1}{2}$  in accordance with the illustration in Fig. 2b. At the end of this section we come back to the case  $p < \frac{1}{2}$ .

Emmy proceeds with three steps made up of quantum thermodynamic primitives with known work and heat contributions [25], Fig. 2b:

$$(\rho, H) \xrightarrow{1} (\rho_1, H^{(1)}) \xrightarrow{2} (\eta^H, H^{(2)}) \xrightarrow{3} (\eta^H, H).$$

In the first step,  $(\rho, H) \xrightarrow{1} (\rho_1, H^{(1)})$ , Emmy isolates the spin from the bath and rotates the B-field such that the variation of the field induces a unitary transformation of the spin into the energy eigenbasis, with unitary  $V = |e_0\rangle\langle 0| + |e_1\rangle\langle 1|$ . The state after this step is

$$\rho_1 = V \rho V^\dagger = a \Pi_0^H + (1-a) \Pi_1^H. \quad (\text{B6})$$

The B-field after this step,  $\vec{B}^{(1)}$ , is chosen such that the new Hamiltonian  $H^{(1)} = -E^{(1)} (\Pi_0^H - \Pi_1^H)$  has eigenvalues  $E^{(1)} = |\vec{\mu}| |\vec{B}^{(1)}| = \frac{k_B T}{2} \ln \frac{a}{1-a}$ , where  $k_B$  is the Boltzmann constant and  $T$  is the temperature of the heat bath that Emmy will use in the next step. This choice of the B-field makes the state  $\rho_1$  a thermal state with respect to  $H^{(1)}$  at temperature  $T$ , i.e.  $\rho_1 = \frac{e^{-\beta H^{(1)}}}{Z^{(1)}}$  with  $Z^{(1)} = \text{tr}[e^{-\beta H^{(1)}}]$  and inverse temperature  $\beta = \frac{1}{k_B T}$ . Since the system was isolated in the first step no heat exchange was possible and the entire average energy change of the system is drawn from the system as work  $\langle W^{(1)} \rangle = -\text{tr}[\rho_1 H^{(1)} - \rho H]$ . Physical constraints may make this process difficult to realise, for instance, pure initial states would require a B-field,  $B^{(1)}$ , of infinite magnitude because thermal states at any finite temperature are only pure if the energy gap is infinite. In this case there is a trade-off between the maximal magnitude the B-field can reach and the precision with which the process is carried out. In the following we assume that the maximal B-field is large enough to make the error in the precision negligibly small.

In the second step,  $(\rho_1, H^{(1)}) \xrightarrow{2} (\eta^H, H^{(2)})$ , Emmy brings the spin in contact with the bath at temperature  $T$ , not affecting the spin's state as it is already thermal. She then quasi-statically decreases the magnitude of the B-field, while keeping the system in contact with the bath at all times, such that the final Hamiltonian is  $H^{(2)} = -E^{(2)} (\Pi_0^H - \Pi_1^H)$  where the B-field is chosen such that  $E^{(2)} = \frac{k_B T}{2} \ln \frac{p}{1-p}$  where  $p$  is the probability of measuring  $-E$  in the initial state,  $\rho$ . The quasi-static evolution means that the system is thermalised at all times, arriving in the final state

$$\frac{e^{-\beta H^{(2)}}}{Z^{(2)}} = p \Pi_0^H + (1-p) \Pi_1^H = \eta^H \quad (\text{B7})$$

which is thermal with respect to  $H^{(2)}$  where  $Z^{(2)} = \text{tr}[e^{-\beta H^{(2)}}]$ . This state is exactly  $\eta^H$ , the desired final state after the projection. The quasi-static process considered here has a known average work given by the free energy difference [25, 29, 39],  $\langle W^{(2)} \rangle = -(F_T^{(2)}(\eta^H) - F_T^{(1)}(\rho_1))$  where  $F_T^{(2)}(\eta^H) = \text{tr}[\eta^H H^{(2)}] - k_B T S(\eta^H)$  and  $F_T^{(1)}(\rho_1) = \text{tr}[\rho_1 H^{(1)}] - k_B T S(\rho_1)$  are standard thermal equilibrium free energies and  $S$  is the von Neumann entropy defined by  $S(\eta^H) = -\text{tr}[\eta^H \log \eta^H]$  and likewise for  $\rho_1$ .

Finally, in the third step,  $(\eta^H, H^{(2)}) \xrightarrow{3} (\eta^H, H)$ , Emmy isolates the spin from the bath and changes the energy levels of the Hamiltonian such that it becomes the initial Hamiltonian  $H$  again. This step is done quickly so that the state of the spin does not change. Because the system is isolated the energy change in this step is entirely due to work  $\langle W^{(3)} \rangle = -\text{tr}[\eta^H (H - H^{(2)})]$ . In total, this thermodynamic process has brought the spin from the quantum state  $(\rho, H)$  to the state  $(\eta^H, H)$  while not changing the energy of the spin,  $\text{tr}[(\rho - \eta^H) H] = 0$ . The overall average work drawn from the spin is

$$\begin{aligned} \langle W \rangle &= \langle W^{(1)} \rangle + \langle W^{(2)} \rangle + \langle W^{(3)} \rangle \\ &= -\text{tr}[\rho_1 H^{(1)}] + \text{tr}[\rho H] - \text{tr}[\eta^H H^{(2)}] + k_B T S(\eta^H) \\ &\quad + \text{tr}[\rho_1 H^{(1)}] - k_B T S(\rho_1) - \text{tr}[\eta^H (H - H^{(2)})] \\ &= k_B T (S(\eta^H) - S(\rho_1)) \\ &= k_B T (S(\eta^H) - S(\rho)) \equiv \langle W_{\text{opt}} \rangle, \end{aligned}$$

showing the optimality of the three step process for the spin example, cf. Eq. (1).

The above example assumed  $p \geq \frac{1}{2}$ . Suppose now that the probability to find the final state  $\eta^H$  in the ground state  $|e_0\rangle$  with respect to the Hamiltonian  $H$  was smaller than to find it in the excited state  $|e_1\rangle$ , i.e.  $p < \frac{1}{2}$ . Proceeding through the three steps described one finds that the mathematics is exactly the same. In particular, after Step 2  $\eta^H$  is a thermal state with respect to  $H^{(2)}$  at inverse temperature  $\beta$ . The only difference occurs in the interpretation as for the Hamiltonian  $H^{(2)}$  the ground state is  $|e_1\rangle$  because  $E^{(2)} = \frac{k_B T}{2} \ln \frac{p}{1-p} < 0$  is negative. This is feasible by making the B-field  $B^{(2)}$  negative, thus swapping the ground and the excited state. Consequently the analysis above and the resulting expression of the total extracted work remain the same.

The work extracted in the individual steps of the thermodynamic projection process can be either positive or negative, depending on the initial state  $\rho$ , the Hamiltonian  $H$  and the temperature  $T$  of the heat bath. Their sum,  $\langle W \rangle$ , is strictly positive whenever the initial state was not diagonal in the energy eigenbasis, a consequence of the entropy increase [17] from  $\rho$  to  $\eta^H$ . On the other hand for classical states – all diagonal in the energy basis – the optimal work for such a projection is always zero.

Appendix D extends the optimality proof of the above three step process to the general finite-dimensional case.

**A note on optimal work extraction at constant average energy.** Assume we are given an initial state  $\rho$  and a non-degenerate Hamiltonian  $H$  for a quantum system. The goal is to find the maximal work that can be obtained in a thermodynamic process that involves a heat bath at temperature  $T$  under the restriction that the average energy of the system after the process is the same as it was before the process,  $U := \text{tr}[\rho H]$ . Using Eq. (2) together with the condition that internal energy does not change this amounts to finding the maximum over the set of states  $\sigma$  with  $\text{tr}[\sigma H] = U$ ,

$$\begin{aligned} \langle W^{\text{max}} \rangle &:= \max_{\sigma} k_B T (S(\sigma) - S(\rho)) \\ &= k_B T (\max_{\sigma} S(\sigma) - S(\rho)). \end{aligned} \quad (\text{B8})$$

It is well-known that at a fixed expectation value of an observable  $H$  the Gibbs states  $\sigma_{\lambda} = e^{-\lambda H} / \text{tr}[e^{-\lambda H}]$  are the states of maximal entropy [40, 41]. Here the parameter  $\lambda$  has to be chosen such that the energy of the Gibbs state matches  $U$  – therefore there is only one  $\sigma_{\lambda^*}$ , with  $\lambda^*$  such that  $\text{tr}[\sigma_{\lambda^*} H] \equiv U$ , that gives the maximum here. The maximum entropy is then

$$\begin{aligned} S(\sigma_{\lambda^*}) &= -\text{tr}[\sigma_{\lambda^*} \ln \sigma_{\lambda^*}] = -\text{tr}[\sigma_{\lambda^*} (-\lambda^* H - \ln \text{tr}[e^{-\lambda^* H}])] \\ &= \lambda^* U + \ln \text{tr}[e^{-\lambda^* H}] \end{aligned} \quad (\text{B9})$$

and the maximum average work that can be extracted from  $\rho$  at fixed average energy  $U$  is then

$$\langle W^{\text{max}} \rangle = k_B T (\lambda^* U + \ln \text{tr}[e^{-\lambda^* H}] - S(\rho)), \quad (\text{B10})$$

For the special case that the system is a qubit (two-dimensional) the optimum Gibbs state for work extraction  $\sigma_{\lambda^*}$  is identical to the projected state  $\eta^H = \sum_{k=0,1} \Pi^{(k)} \rho \Pi^{(k)}$  and the maximal work that can be drawn from a system starting in state  $\rho$ , while keeping its average energy fixed, is  $\langle W_{\text{opt}} \rangle$  in Eq. (1). To see this we expand  $H = E^{(0)} \Pi^{(0)} + E^{(1)} \Pi^{(1)}$  and  $U = p E^{(0)} + (1-p) E^{(1)}$  with  $p := \text{tr}[\rho \Pi^{(0)}]$ . Now here  $\lambda^*$  must be chosen such that  $\sigma_{\lambda^*} = e^{-\lambda^* H} / \text{tr}[e^{-\lambda^* H}] = p \Pi^{(0)} + (1-p) \Pi^{(1)}$ , i.e.  $p = e^{-\lambda^* E^{(0)}} / (e^{-\lambda^* E^{(0)}} + e^{-\lambda^* E^{(1)}})$ , so that  $\sigma_{\lambda^*}$  has just the right energy  $\text{tr}[\sigma_{\lambda^*} H] = U$ . On the other hand the projection state has the same expansion,  $\eta^H = p \Pi^{(0)} + (1-p) \Pi^{(1)} = \sigma_{\lambda^*}$ . We note that this coincidence is not true for higher dimensional systems where the energy-projected state  $\eta^H$  will in general have a non-monotonous, non-canonical distribution in its energy eigenbasis, while  $\sigma_{\lambda^*}$  must be Gibbs-distributed.

Considering the illustration in Fig. 2b, the qubit states  $\sigma$  fulfilling the condition  $\text{tr}[\sigma H]$  are located on the plane which contains  $\rho$  and is perpendicular to the  $|e_0\rangle$ - $|e_1\rangle$ -axis. On the other hand, in the Bloch picture a state has higher entropy the closer it is to the center of the sphere. Hence, the optimal final state when extracting



work from  $\rho$  while conserving the average energy of the system is the state  $\rho$  projected to the  $|e_0\rangle$ - $|e_1\rangle$ -axis, i.e.  $\eta^H$ .

### Appendix C: Work storage system

In the previous section it was stated that work can be drawn from a quantum system when undergoing a thermodynamic projection process. But where has the work gone to?

There are two approaches of accounting for work that are mirror images to each other. One approach [6, 22, 24–26, 29, 39, 42, 43] focusses on the work that the system exchanges, as described above. Here it is often not explicitly mentioned where the work goes to, but the only place it can go to are the externally controlled energy sources, see Fig. 1. Another way of accounting is to explicitly introduce a work system to store the work drawn [23, 30]. One way of doing so in an average scenario is to introduce [23] a ‘suspended weight on a string’, described by a quantum system  $W$ , that could be raised or lowered to store work or draw work from it. Specifically, the Hamiltonian of the work storage system is defined as  $H^W = m g x$ , representing the energy of a weight of mass  $m$  in the gravitational field with acceleration  $g$  at height  $x$ . In addition, an explicit thermal bath  $B$  is introduced [40, 41] consisting of a separate quantum system in a thermal (or Gibbs) state  $\tau^B$ . Both, the explicit work storage system and the heat bath are illustrated in Fig. 1. In the latter approach the total system starts in a product state of system  $S$  (e.g. spin), bath  $B$ , and weight  $W$ ,  $\rho^{SBW} = \rho^S \otimes \tau^B \otimes \omega^W$ , which together undergo *average energy conserving unitary evolution* with  $V$ :

$$\rho^{SBW} \mapsto \sigma^{SBW} = V \rho^{SBW} V^\dagger. \quad (C1)$$

The assumption is that the total Hamiltonian is the sum of local terms,  $H^{SBW} = H^S + H^B + H^W$ . The average energy conservation constraint then reads  $\text{tr}[(\sigma^{SBW} - \rho^{SBW}) H^{SBW}] = 0$  and the average work extracted to the work storage system,  $\langle W \rangle$ , is identified with

$$\text{tr}[(\sigma^{SBW} - \rho^{SBW}) H^W] = \langle W \rangle. \quad (C2)$$

Both the implicit and the explicit treatment of work are equivalent in the sense that the results obtained in one language can be translated in the other and vice versa. In particular, the implicit description used in this text [25] has an equivalent explicit formulation [23].

In the next section we will discuss single-shot extractable work in a projection process. One possibility to define work in this context is to chose the explicit work storage system as a ‘work qubit’ with a specific energy gap which has to be in a pure energy eigenstate before and after the protocol [30]. This way it is guaranteed that full knowledge about its state is present at all times and the work is stored in an ordered form. In this scenario the allowed unitary operations  $V$  on the whole system

$SBW$  have to conserve the energy exactly, not only on average, which amounts to  $[V, H^{SBW}] = 0$ .

### Appendix D: General three step process

It is straightforward to generalise the proof of optimality from the two-dimensional spin-1/2 example to thermodynamic projection processes in dimension  $d$ . Again the projectors  $\{\Pi_k^H\}_k$  map onto the energy eigenspaces of the Hamiltonian,  $H = \sum_k E_k^{(0)} \Pi_k^H$ , where  $E_k^{(0)}$ ,  $k = 1, \dots, d$ , are the energy eigenvalues. A general initial state can be written as  $\rho = \sum_{j=1}^d a_j |j\rangle\langle j|$  where  $a_j \geq 0$  are probabilities,  $\sum_{j=1}^d a_j = 1$ ,  $|j\rangle\langle j|$  are rank-1 projectors on the corresponding eigenvectors  $|j\rangle$ , and  $j = 1, \dots, d$ . A unitary operation,  $V$ , is now chosen such that it brings the initial configuration  $(\rho, H)$  into the new diagonal and thermal configuration  $(\rho_1 = V \rho V^\dagger, H^{(1)})$  where  $\rho_1 = \sum_k a_k \Pi_k^H$  and  $H^{(1)} = \sum_k E_k^{(1)} \Pi_k^H$ . The new energy eigenvalues,  $E_k^{(1)}$ , are adjusted such that the probabilities  $a_k$  are thermally distributed with respect to  $H^{(1)}$  for the bath temperature  $T$ . Adjusting the Hamiltonian eigenvalues while letting the state thermalise at all times now results in a isothermal quasi-static operation from  $(\rho_1, H^{(1)})$  to  $(\eta^H = \sum_k p_k \Pi_k^H, H^{(2)} = \sum_k E_k^{(2)} \Pi_k^H)$ . Here the new energy eigenvalues,  $E_k^{(2)}$ , are chosen to be thermal (at  $T$ ) for the state’s probabilities which are given by  $p_k = \text{tr}[\rho \Pi_k^H]$ . Finally, a quench brings the thermal configuration  $(\eta^H, H^{(2)})$  quickly into the non-equilibrium state  $(\eta^H, H)$ . The average work for this overall process is  $\langle W \rangle = \sum_{j=1}^3 \langle W^{(j)} \rangle$  where  $\langle W^{(1)} \rangle = -\text{tr}[H^{(1)} \rho_1 - H \rho]$  and  $\langle W^{(3)} \rangle = -\text{tr}[H \eta^H - H^{(2)} \eta^H]$  because the first and third steps are unitary (Q0+T0). The quasistatic step’s work is [25, 29]  $\langle W^{(2)} \rangle = -F_T^{(2)} + F_T^{(1)}$  where  $F_T^{(1)} = \text{tr}[H^{(1)} \rho_1] - k_B T S(\rho_1)$  is the thermal equilibrium free energy for Hamiltonian  $H^{(1)}$ , and similarly,  $F_T^{(2)} = \text{tr}[H^{(2)} \eta^H] - k_B T S(\eta^H)$ . Summing up and using  $\text{tr}[H(\rho - \eta^H)] = 0$ , one obtains

$$\langle W \rangle = k_B T (S(\eta^H) - S(\rho)) \equiv \langle W_{\text{opt}} \rangle, \quad (D1)$$

concluding the optimality proof of the process sequence.

### Appendix E: Lower bound on entropy change

The entropy change during a projection with projectors  $\{\Pi_k^P = |\phi_k\rangle\langle\phi_k|\}_k$  can be lower bounded. In the following,  $\|B\|_2 = \sqrt{\text{tr}[B^\dagger B]}$  denotes the Hilbert-Schmidt norm of a linear operator  $B$  acting on a  $d$ -dimensional Hilbert space describing the quantum system of interest. The lower bound reads [35]

$$\Delta S^P := S(\eta^P) - S(\rho) \geq \frac{1}{2} \left\| \rho - \frac{\mathbb{1}}{d} \right\|_2^2 \Delta A^P. \quad (E1)$$

Here,  $S$  is the von Neumann entropy,  $\rho$  the initial state and  $\eta^P = \sum_k \Pi_k^P \rho \Pi_k^P$  the final state after the projection process. Furthermore,  $\Delta A^P$  is the second smallest eigenvalue of the matrix  $\mathbb{1} - M^T M$  where  $M$  is the doubly stochastic matrix given by the entries  $M_{kl} = |\langle \phi_k | l \rangle|^2$  and  $\{|l\rangle\}_l$  is the eigenbasis of the initial state  $\rho$ .

Considering the two main terms on the right hand side of Eq. (E1) separately,  $\|\rho - \mathbb{1}/d\|_2^2$  and  $\Delta A^P$ , it becomes apparent that they characterise different properties of the initial state. The first term measures the distance of  $\rho$  to the fully mixed state,  $\mathbb{1}/d$ , and quantifies the purity of  $\rho$ . It is maximal for all pure initial states and zero if and only if  $\rho = \mathbb{1}/d$ . In the special case of a spin-1/2 system it can be directly related to the length of the Bloch vector describing  $\rho$  in the Bloch representation, a link that will be established below. The second term,  $\Delta A^P$ , is related to the overlap of the eigenbasis of  $\rho$ ,  $\{|l\rangle\}_l$ , and the projective basis,  $\{|\phi_k\rangle\}_k$ . It is zero if they are the same and maximal if they are mutually unbiased [36, 37]. This can be seen as follows: if the two bases are the same, then the matrix  $M$  is a permutation and consequently  $M^T M$  is the identity. In this case,  $\mathbb{1} - M^T M$  is the zero matrix and thus  $\Delta A^P = 0$ . If  $\{|\phi_k\rangle\}_k$  and  $\{|l\rangle\}_l$  are mutually unbiased, i.e. if they fulfil  $|\langle \phi_k | l \rangle|^2 = 1/d$  for all  $k, l$ , the matrix  $M$  and thus also  $M^T M$  is a rank-1 projector onto the space spanned by the vector  $(1, \dots, 1)^T$ . Hence,  $\mathbb{1} - M^T M$  has eigenvalues  $\{0, 1, \dots, 1\}$ . One finds that the second largest eigenvalue is  $\Delta A^P = 1$ , which is also the maximal eigenvalue the matrix  $\mathbb{1} - M^T M$  can have [38].

In the special case of the spin-1/2 system shown in Fig. 2b, the bound reads  $\Delta S^H(\rho) \geq \frac{1}{4} |\vec{s}_\rho|^2 \sin^2 \theta$ , where  $\vec{s}_\rho$  is the Bloch vector of the initial state and  $\theta$  is the angle between the eigenbasis of  $\rho$ ,  $\{|0\rangle, |1\rangle\}$ , and the projective energy basis,  $\{|e_0\rangle, |e_1\rangle\}$ . Let  $\rho = a|0\rangle\langle 0| + (1-a)|1\rangle\langle 1|$  be the initial state of the qubit. Furthermore, let  $\eta^H = p|e_0\rangle\langle e_0| + (1-p)|e_1\rangle\langle e_1|$  be the final state after the energy projection, where  $p = \text{tr}[|e_0\rangle\langle e_0| \rho]$  is the probability to obtain  $|e_0\rangle$ . As argued in Appendix B w.l.o.g. we can assume that  $a \geq \frac{1}{2}, p \geq \frac{1}{2}$ . In the Bloch representation one can write  $\rho = \frac{1}{2}(\mathbb{1} + \vec{s} \cdot \vec{\sigma})$  and  $\eta^H = \frac{1}{2}(\mathbb{1} + \vec{t} \cdot \vec{\sigma})$ . Here we used a different notation for the Bloch vectors of  $\rho$ ,  $\vec{s} := \vec{s}_\rho$ , and  $\eta^H$ ,  $\vec{t} := \vec{s}_\eta$ , for readability. The Pauli matrices are self-adjoint and fulfil  $\text{tr}[\sigma_i \sigma_j] = 2\delta_{ij}$ . Hence we find

$$\begin{aligned} \left\| \rho - \frac{\mathbb{1}}{2} \right\|_2^2 &= \text{tr} \left[ \left( \frac{1}{2} \vec{s} \cdot \vec{\sigma} \right) \left( \frac{1}{2} \vec{s} \cdot \vec{\sigma} \right) \right] \\ &= \sum_{i,j=1}^3 \frac{1}{4} s_i s_j \text{tr}[\sigma_i \sigma_j] = \frac{1}{2} |\vec{s}|^2, \end{aligned} \quad (\text{E2})$$

where  $|\cdot|$  is the Euclidean metric in  $\mathbb{R}^3$ . This proves the form of the first factor in the bound. For the factor  $\Delta A^H$  notice that by assumption  $a \geq \frac{1}{2}, p \geq \frac{1}{2}$  and thus we can write  $|e_0\rangle\langle e_0| = \frac{1}{2}(\mathbb{1} + \frac{\vec{t}}{|\vec{t}|} \cdot \vec{\sigma})$  and  $|0\rangle\langle 0| = \frac{1}{2}(\mathbb{1} + \frac{\vec{s}}{|\vec{s}|} \cdot \vec{\sigma})$ .

Therefore

$$\begin{aligned} |\langle e_0 | 0 \rangle|^2 &= \text{tr}[|0\rangle\langle 0| |e_0\rangle\langle e_0|] \\ &= \frac{1}{4} \text{tr} \left[ \mathbb{1} + \frac{s_i}{|\vec{s}|} \sigma_i + \frac{t_i}{|\vec{t}|} \sigma_i + \frac{s_i t_j}{|\vec{s}| |\vec{t}|} \sigma_i \sigma_j \right] \\ &= \frac{1}{4} \left[ 2 + 0 + 0 + 2 \frac{\vec{s} \cdot \vec{t}}{|\vec{s}| |\vec{t}|} \right] = \frac{1}{2} [1 + \cos \theta], \end{aligned} \quad (\text{E3})$$

where summations over double indices are assumed and  $\theta$  is the angle between the two Bloch vectors of  $\rho$  and  $\eta^H$ , see Fig. 2b. By using orthonormality of both bases  $\{|l\rangle\}_l$  and  $\{|e_k\rangle\}_k$  it is shown that in the qubit case the matrices  $M$  and  $M^T M$  have the form

$$M = \begin{pmatrix} |\langle e_0 | 0 \rangle|^2 & |\langle e_0 | 1 \rangle|^2 \\ |\langle e_1 | 0 \rangle|^2 & |\langle e_1 | 1 \rangle|^2 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + \cos \theta & 1 - \cos \theta \\ 1 - \cos \theta & 1 + \cos \theta \end{pmatrix} \quad (\text{E4})$$

and

$$M^T M = \frac{1}{2} \begin{pmatrix} 1 + \cos^2 \theta & 1 - \cos^2 \theta \\ 1 - \cos^2 \theta & 1 + \cos^2 \theta \end{pmatrix}. \quad (\text{E5})$$

Computing the eigenvalues  $\{\lambda_1, \lambda_2\}$  of  $M^T M$  yields  $\lambda_1 = 1$  and  $\lambda_2 = \cos^2 \theta$ . Therefore  $\Delta A^H = 1 - \lambda_2 = 1 - \cos^2 \theta = \sin^2 \theta$ . In total, this concludes the proof of the bound for the special case of a spin-1/2 system when projected in the energy eigenbasis  $\{\Pi_k^H\}_k = \{|e_0\rangle, |e_1\rangle\}$ ,

$$\Delta S^H(\rho) \geq \frac{1}{4} |\vec{s}_\rho|^2 \sin^2 \theta. \quad (\text{E6})$$

To further illustrate the bound consider the special case when the initial state  $\rho$  is pure and its eigenbasis mutually unbiased with respect to the energy eigenbasis,  $\{|e_0\rangle, |e_1\rangle\}$ . In this case the final state after the projection,  $\eta^H$ , is maximally mixed and we find

$$\Delta S^H(\rho) = S(\eta^H) - S(\rho) = \ln 2 - 0 \approx 0.69. \quad (\text{E7})$$

Here, the lower bound is equal to  $\frac{1}{4} = 0.25$  because  $|\vec{s}_\rho| = 1$  for a pure state  $\rho$  and  $\sin^2 \theta = 1$  for mutually unbiased bases. Thus in this example the bound is not particularly tight.

## Appendix F: Single-shot analysis

Instead of performing a thermodynamic process on an ensemble of  $N$  identical and independent copies one can consider a single run of the process. Two major recent frameworks [29, 30] have been developed to describe the optimal work that can be drawn from a system in a single run. The proposal by Åberg [29], involves changes of the Hamiltonian and identifies work with the deterministic energy change of the system when undergoing a unitary process. The proposal by Horodecki-Oppenheim [30], is formulated in terms of *thermal operations* [44], where work is associated with raising a two-level system,

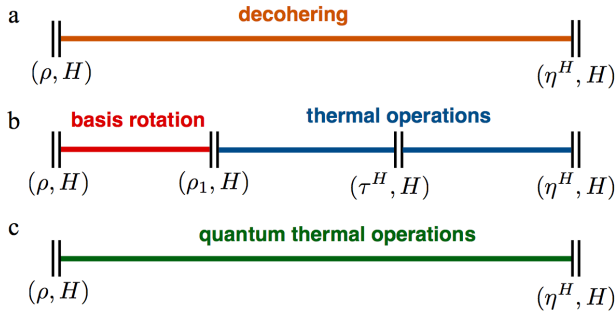


FIG. 4: **Illustration of possible quantum thermodynamic processes that transform  $(\rho, H)$  to  $(\eta^H, H)$ .** **a.** Decohering the state in the energy basis extracts no work. **b.** To perform a consistency check between the average and single-shot results it is possible to split the process into a basis rotation to  $(\rho_1, H)$  with unknown single-shot work, but known average work, and two thermal operations that pass through the thermal state  $(\tau^H, H)$  and are treatable in the single-shot framework [30]. **c.** General quantum thermodynamic processes could allow coherences and need not pass through intermediate fixed states.

called the ‘work qubit’, with energy gap  $W$  deterministically from the ground to the excited state.

However, when attempting to apply these two frameworks to find the single-shot work for the energy projections  $\rho \rightarrow \eta^H$  captured by Eq. (1) one encounters an obstacle: both frameworks only apply to processes between initial and final states that are classical, i.e. states that are diagonal in the energy basis. Åberg discusses coherences in a separate framework [12], which does however not cover single-shot work extraction and only focusses on average quantities, similar to those in other references [23, 25]. Horodecki-Oppenheim suggest that quantum states with coherences with respect to the energy eigenbasis are first decohered before applying the single-shot protocol. As discussed, apart from decohering there are other thermodynamic projection processes that map the initial state with coherences,  $\rho$ , to the final state  $\eta^H = \sum_k \Pi_k^H \rho \Pi_k^H$  without coherences, where  $\Pi_k^H$  are the projectors on the energy eigenstates of the Hamiltonian,  $H$ . Eq. (1) shows that the average work extracted in an optimal thermodynamic projection process is strictly positive while the decoherence process has zero work. Therefore one may expect a positive optimal work for projections also in the single-shot setting, with decohering a suboptimal choice, see Fig. 4.

Since our focus here is the  $N \rightarrow \infty$  limit we will not aim to construct the single-shot case. Instead, to establish a notion of consistency between the average analysis and previous single-shot work results we consider the sequence  $(\rho, H) \xrightarrow{a} (\rho_1, H) \xrightarrow{b} (\tau^H, H) \xrightarrow{c} (\eta^H, H)$  in which the Hamiltonian before and after each step are the same and  $\rho_1$  is the rotated state defined above, Eq. (B6). Here  $\tau^H = \frac{e^{-\beta H}}{Z}$  is the thermal state for Hamiltonian  $H$  at inverse temperature  $\beta = \frac{1}{k_B T}$  and  $Z = \text{tr}[e^{-\beta H}]$  is the partition function. Step *a* of this sequence rotates the initial non-diagonal state  $\rho$  to the diagonal

state  $\rho_1$ . As discussed, it cannot be treated with the single-shot framework [29, 30] but it is possible to associate an average extracted work with this unitary process,  $\langle W^{(a)} \rangle = \text{tr}[(\rho - \rho_1) H]$ . A single-shot analysis according to Horodecki-Oppenheim [30] can then be performed for the diagonal steps *b* and *c*. This is possible because the steps go via the thermal state  $\tau^H$ . Step *b* brings  $\rho_1$  to  $\tau^H$  and allows the extraction of the single-shot work [30]

$$W^{(b)} = k_B T \ln 2 D_{\min}^{\varepsilon}(\rho_1 || \tau^H), \quad (\text{F1})$$

where  $D_{\min}^{\varepsilon}$  is the smooth min-relative entropy [45] and  $\varepsilon \geq 0$  is the allowed failure probability of the process. Similarly, in Step *c* the final state  $\eta^H$  is formed from the thermal state by applying a protocol that costs work. This work is [30]

$$W^{(c)} = -k_B T \ln 2 D_{\max}^{\varepsilon}(\eta^H || \tau^H), \quad (\text{F2})$$

where  $D_{\max}^{\varepsilon}$  is the smooth max-relative entropy [45]. In total, the single-shot work associated to Steps *b* and *c* of the process is  $W^{(b)} + W^{(c)} = k_B T \ln 2 (D_{\min}^{\varepsilon}(\rho_1 || \tau^H) - D_{\max}^{\varepsilon}(\eta^H || \tau^H))$  with failure probability at most  $2\varepsilon - \varepsilon^2 \approx 2\varepsilon$ , when  $\varepsilon$  is small.

To show consistency we now consider the average expected work extracted per copy if the single-shot protocol is carried out on  $N \rightarrow \infty$  i.i.d. copies of the system. In such a calculation the work computed is an average value which is why  $\langle W^{(a)} \rangle$ , the average work contribution of the basis rotation in Step *a*, can be taken into account too. One obtains a total average work per copy of

$$\begin{aligned} \langle W \rangle &= \langle W^{(a)} \rangle + k_B T \ln 2 \lim_{\varepsilon \rightarrow 0} \lim_{N \rightarrow \infty} \frac{1}{N} \\ &\quad [D_{\min}^{\varepsilon}(\rho_1^{\otimes N} || (\tau^H)^{\otimes N}) - D_{\max}^{\varepsilon}((\eta^H)^{\otimes N} || (\tau^H)^{\otimes N})] \\ &= \langle W^{(a)} \rangle + k_B T \ln 2 [D(\rho_1 || \tau^H) - D(\eta^H || \tau^H)] \\ &= \langle W^{(a)} \rangle + \text{tr}[\rho_1 H] + \ln Z - k_B T S(\rho_1) \\ &\quad - \text{tr}[\eta^H H] - \ln Z + k_B T S(\eta^H) \\ &= k_B T (S(\eta^H) - S(\rho)) \\ &\equiv \langle W_{\text{opt}} \rangle, \end{aligned} \quad (\text{F3})$$

where we have used the quantum asymptotic equipartition theorem for relative entropies [46, 47] in the second line.  $D(\cdot || \cdot)$  is the standard quantum relative entropy defined by  $D(\eta^H || \tau^H) = \text{tr}[\eta^H (\log \eta^H - \log \tau^H)]$  and likewise for  $\rho_1$ , where  $\log$  is the logarithm to base 2. The quantities  $D_{\min}^{\varepsilon}$  and  $D_{\max}^{\varepsilon}$  as well as their regularized version, the standard quantum relative entropy  $D$ , can be seen as different measures characterizing the distance between two states. When applied here, they measure the ‘distance’ between the thermal state  $\tau^H$  and another diagonal state in such a way that the operational meaning of this distance is given by the work one has to invest or is able to extract when transforming one into the other.

The derivation shows that in the asymptotic limit the optimal average work is recovered from the single-shot

components. But it is important to realise that from Eq. (F3) one cannot conclude that the above single-shot process forming  $\eta^H$  from  $\rho_1$  is optimal. Going via the thermal state is just one option which is particularly convenient in this case as the processes of maximal work extraction and work of formation from the thermal state have been treated in the single-shot scenario [30]. It is an open question whether there are better single-shot protocols for general thermodynamic transformations, see Fig. 4b & 4c. The introduction of “catalysts” in single-shot thermodynamics [8] provides a promising avenue to establish bounds on the work that can be drawn from a state with coherences during a projection in the single-shot setting.

After making public our results on average work associated with removing coherences in thermodynamic projection processes very recently a paper appeared [31] that derives the work that can be extracted when removing coherences in a single-shot setting. In this paper the previously mentioned framework describing the catalytic role of coherence in thermodynamics by Åberg [12] is used together with insights from reference frames in quantum information theory. These results are in agreement with our findings and strengthen our conclusion that coherences are a fundamental feature distinguishing quantum from classical thermodynamics.

### Appendix G: Quantum work fluctuation relation

A common route of deriving the quantum Jarzynski equation is as follows [5, 6, 42]. A quantum system is initialised in a thermal state  $\rho_0 = \sum_n e^{-\beta(E_n^{(0)} - F_T^{(0)})} \Pi_n^{(0)}$  for a given Hamiltonian  $H^{(0)} = \sum_n E_n^{(0)} \Pi_n^{(0)}$ , with energy eigenvector projectors  $\Pi_n^{(0)} = |e_n^{(0)}\rangle\langle e_n^{(0)}|$ , at given inverse temperature  $\beta = 1/(k_B T)$ . Here  $F_T^{(0)} = -k_B T \ln(\sum_n e^{-\beta E_n^{(0)}})$  is the initial free energy associated with the initial Hamiltonian  $H^{(0)}$ . The aim is to calculate the average exponentiated work,  $\langle e^{\beta W} \rangle$ , that the quantum system will exchange when undergoing a unitary process  $V$  that is generated by varying the Hamiltonian in time, i.e.  $H^{(t)}$ , from  $H^{(0)}$  to a final  $H^{(\tau)} = \sum_m E_m^{(\tau)} \Pi_m^{(\tau)}$ . The final state after the unitary is the non-equilibrium state  $\rho_\tau := V \rho_0 V^\dagger$ , see Fig. 3a.

To identify the work for an individual run of the experiment the energy of the system is measured at the beginning, by projecting into  $\Pi_n^{(0)}$ , and at the end, by projecting in the final energy basis  $\Pi_m^{(\tau)} = |e_m^{(\tau)}\rangle\langle e_m^{(\tau)}|$ . The (extracted) fluctuating work  $W$  identified with each transition  $|e_n^{(0)}\rangle \rightarrow |e_m^{(\tau)}\rangle$  is the (negative) observed *fluctuating energy difference* of the system

$$\Delta E = E_m^{(\tau)} - E_n^{(0)} = -W. \quad (\text{G1})$$

The average exponentiated work then becomes

$$\langle e^{\beta W} \rangle = \sum_{m,n} e^{-\beta(E_m^{(\tau)} - E_n^{(0)})} p_{m,n}^\tau, \quad (\text{G2})$$

where  $p_{m,n}^\tau$  are the transition probabilities for energy jumps starting in  $|e_n^{(0)}\rangle$  and ending in  $|e_m^{(\tau)}\rangle$  at time  $\tau$ . These probabilities are given by

$$\begin{aligned} p_{m,n}^\tau &= \text{tr}[\Pi_m^{(\tau)} V \Pi_n^{(0)} \rho_0 \Pi_n^{(0)} V^\dagger \Pi_m^{(\tau)}] \\ &= e^{-\beta(E_n^{(0)} - F_T^{(0)})} \text{tr}[\Pi_m^{(\tau)} V \Pi_n^{(0)} V^\dagger], \end{aligned} \quad (\text{G3})$$

simplifying the exponentiated average work to

$$\langle e^{\beta W} \rangle = e^{\beta F_T^{(0)}} \sum_m e^{-\beta E_m^{(\tau)}} \sum_n \text{tr}[\Pi_m^{(\tau)} V \Pi_n^{(0)} V^\dagger]. \quad (\text{G4})$$

The completeness of the projectors,  $\sum_n \Pi_n^{(0)} = \mathbb{1}$ , now finally results in the well-known quantum Jarzynski work relation

$$\langle e^{\beta W} \rangle = e^{\beta F_T^{(0)}} \sum_m e^{-\beta E_m^{(\tau)}} = e^{-\beta \Delta F}, \quad (\text{G5})$$

where  $\Delta F = F_T^{(\tau)} - F_T^{(0)}$  is the difference of the equilibrium free energies corresponding to the final and initial Hamiltonians, i.e.  $F_T^{(\tau)} = -k_B T \ln(\sum_m e^{-\beta E_m^{(\tau)}})$ . Similarly, the average work extracted from the system is the average energy difference between  $\rho_0$  and  $\rho_\tau$

$$\begin{aligned} \langle W^{\text{unitary}} \rangle &= - \sum_{m,n} (E_m^{(\tau)} - E_n^{(0)}) p_{m,n}^\tau \\ &= - \sum_m E_m^{(\tau)} p_m^\tau + \sum_n E_n^{(0)} p_n^0 \\ &= -\text{tr}[H^{(\tau)} \rho_\tau] + \text{tr}[H^{(0)} \rho_0] \\ &= -U(\rho_\tau) + U(\rho_0) \end{aligned} \quad (\text{G6})$$

where  $p_m^\tau := \sum_n p_{m,n}^\tau = \text{tr}[\rho_\tau \Pi_m^{(\tau)}]$  are the probabilities to find energies  $E_m^{(\tau)}$  when a measurement of  $H^{(\tau)}$  is performed on  $\rho_\tau$ . By construction, the initial probabilities to find energies  $E_n^{(0)}$  are just the thermal probabilities,  $p_n^0 := \sum_m p_{m,n}^\tau = \text{tr}[e^{-\beta(E_n^{(0)} - F_T^{(0)})} \Pi_n^{(0)}] = e^{-\beta(E_n^{(0)} - F_T^{(0)})}$ .

The derivations of the average work,  $\langle W^{\text{unitary}} \rangle = U(\rho_0) - U(\rho_\tau)$ , as well as the average exponentiated work, i.e. the quantum Jarzynski equality,  $\langle e^{\beta W} \rangle = e^{-\beta \Delta F}$ , are based on Eq. (G1) which was made assuming that the initial and final state of the process that is being characterised are  $\rho_0$  and  $\rho_\tau$ . There is no question that the mathematical details of the derivations of the above relation are sound. Experimentally, there is however a need to acquire knowledge of the fluctuating energy to quantify the work and this requires the implementation of the second measurement, see Fig. 3b. Only after the measurement has been made can theoretical predictions be

tested. The measurement is an unavoidable non-unitary component of the overall experimental process. Specifically, the ensemble state after the unitary,  $\rho_\tau$ , is further altered by the measurement to result in the final state  $\eta_\tau$ , i.e. it is the state  $\rho_\tau$  with any coherences in the energy basis of  $H^{(\tau)}$  removed.

While the experimentally observed average energy difference is not affected by the measurement step, i.e.  $U(\eta_\tau) - U(\rho_0) = U(\rho_\tau) - U(\rho_0)$ , the entropy difference does change, i.e.  $S(\eta_\tau) - S(\rho_0) \neq S(\rho_\tau) - S(\rho_0) = 0$ . This means that the system may absorb heat,  $\langle Q^{\text{abs}} \rangle$ , during the measurement step, indicated in Fig. 3b. Its actual value depends on *how* the measurement is conducted with the optimal heat positive,  $\langle Q_{\text{opt}}^{\text{abs}} \rangle = k_B T (S(\eta_\tau) - S(\rho_\tau)) \geq 0$ . Since  $\Delta U = \langle Q^{\text{abs}} \rangle - \langle W \rangle$  (T1) this implies that in an experimental implementation of the Jarzynski relation the work done by the system on average can be more than previously thought, with the optimal value being  $\langle W_{\text{opt}} \rangle = \langle W^{\text{unitary}} \rangle + k_B T (S(\eta_\tau) - S(\rho_\tau))$ . In the special case that the average heat  $\langle Q^{\text{abs}} \rangle$  is zero it is possible (although not necessary) that Eq. (G1), and thus the standard Jarzynski expression  $\langle e^{\beta W} \rangle = e^{-\beta \Delta F}$ , are correct. In particular this applies to classical measurements. We conclude that the suitability of identifying  $W = -\Delta E$ , and hence the validity of the quantum Jarzynski work relation depends on the details of the physical process that implements the measurement.

Quantum work fluctuation relations that have only one measurement [48, 49], instead of the two discussed above, offer a feasible route of measuring work fluctuations experimentally. Instead of measuring separately the initial and final fluctuating energies,  $E_n^{(0)}$  and  $E_m^{(\tau)}$ , to establish their joint probabilities, this method acquires *only* knowledge of the joint probabilities by measuring energy differences  $\Delta E$  directly. But also here is one final measurement, in general on a non-diagonal state, needed.

## Appendix H: Access to correlated auxiliary systems

Similarly to erasure with a correlated memory [14] one can consider projections on a system  $S$  that is correlated with an ancilla  $A$  the experimenter has access to. Assuming a total Hamiltonian  $H^{SA} = H^S \otimes \mathbb{1}^A + \mathbb{1}^S \otimes H^A$ , we denote the global initial state by  $\rho^{SA}$  and its marginals on  $S$  and  $A$  by  $\rho^S = \text{tr}_A[\rho^{SA}]$  and  $\rho^A = \text{tr}_S[\rho^{SA}]$ , respectively.

**A note on notation.** For clarity we employ a slightly different notation here. The roles of initial state  $\rho$  and final state  $\eta$  are the same as in the main text and the previous sections of the Appendix. However, now the superscripts of the final state  $\eta$  no longer denote the projection basis but the system for which  $\eta$  describes the state. For instance,  $\eta^S$  denotes the reduced state after the projection on system  $S$  alone. The same holds for the superscript of the initial state,  $\rho^{SA}, \rho^S$  and  $\rho^A$ , and the Hamiltonians  $H^{SA}, H^S$  and  $H^A$ . Only

the superscript  $\mathcal{P}$  of the mutually orthogonal rank-1 projectors  $\{\Pi_k^{\mathcal{P}}\}_k$  acting on system  $S$  is kept to indicate which basis is being projected in.

For an initial global state  $\rho^{SA}$  of system and ancilla a local projection map on  $S$  results in a new global state

$$\eta^{SA} = \sum_k (\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A) \rho^{SA} (\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A). \quad (\text{H1})$$

Due to the properties of the projectors the marginal state on  $A$  is unchanged,

$$\begin{aligned} \eta^A &= \text{tr}_S[\eta^{SA}] = \sum_k \text{tr}_S[(\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A) \rho^{SA} (\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A)] \\ &= \sum_k \text{tr}_S[\rho^{SA} (\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A)] = \text{tr}_S[\rho^{SA}] = \rho^A. \end{aligned} \quad (\text{H2})$$

The reduced state of the system becomes  $\eta^S \equiv \text{tr}_A[\eta^{SA}] = \sum_k \Pi_k^{\mathcal{P}} \rho^S \Pi_k^{\mathcal{P}} = \sum_k p_k \Pi_k^{\mathcal{P}}$  where  $p_k = \text{tr}[\rho^S \Pi_k^{\mathcal{P}}]$ , and the conditional states on  $A$  after the process are denoted  $\eta_k^A = p_k^{-1} \text{tr}_S[(\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A) \rho^{SA} (\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A)]$  for all  $k$ . The global entropy change associated with the local projection is

$$\begin{aligned} \Delta S^{\mathcal{P}} &= S(\eta^{SA}) - S(\rho^{SA}) \\ &= S(\{p_k\}) + \sum_k p_k S(\eta_k^A) - S(\rho^{SA}) \\ &= S(\eta^S) - S(\rho^S) + S(\rho^S) + \sum_k p_k S(\eta_k^A) - S(\rho^{SA}) \\ &= \Delta S^{\mathcal{P}} + \delta^{\mathcal{P}}(A : S). \end{aligned} \quad (\text{H3})$$

In the second equality it was used that  $\eta^{SA}$  is a classical-quantum-state and  $S(\{p_k\}) = -\sum_k p_k \ln p_k$  stands for the classical Shannon entropy [50] which is equal to the von Neumann entropy of  $\eta^S$  because the final state on  $S$  is a classical mixture of states from the projective basis. Here we defined a measure of correlations between the ancilla and the system,  $\delta^{\mathcal{P}}(A : S) = S(\rho^S) - S(\rho^{SA}) + \sum_k p_k S(\eta_k^A)$ , related to the quantum discord. It depends on the projectors  $\{\Pi_k^{\mathcal{P}}\}_k$  and is always positive [32, 51]. Thus the entropy change of  $SA$  can be bigger than the local entropy change,  $\Delta S^{\mathcal{P}} = S(\eta^S) - S(\rho^S)$ , on the system alone.

As shown before, Eq. (2), the optimal extractable work in a thermodynamic projection process on system  $S$  alone is  $\langle W_{\text{opt}} \rangle = k_B T \Delta S^{\mathcal{P}} - \Delta U^{\mathcal{P}}$ , where  $\Delta S^{\mathcal{P}}$  is the entropy change of the system and  $\Delta U^{\mathcal{P}}$  its change in internal energy. This result stays intact when generalizing to projections in the presence of ancillary systems if one takes the total changes of these quantities on  $SA$  instead of the change on  $S$  only. In the global process the total internal energy change is equal to the energy change of the system only as the local state of the ancilla is unchanged and the total Hamiltonian is the sum of local Hamiltonians. Thus using side information the overall optimal

extractable work amounts to

$$\begin{aligned}\langle W_{\text{opt}} \rangle &= k_B T \Delta \mathcal{S}^{\mathcal{P}} - \Delta U^{\mathcal{P}} \\ &= k_B T \Delta \mathcal{S}^{\mathcal{P}} + k_B T \delta^{\mathcal{P}}(A : S) - \Delta U^{\mathcal{P}} \quad (\text{H4}) \\ &= \langle W_{\text{opt}} \rangle + k_B T \delta^{\mathcal{P}}(A : S),\end{aligned}$$

where  $\langle W_{\text{opt}} \rangle$  is the work of an optimal thermodynamic projection process without access to correlated systems, Eq. (2).

Discord was first discussed in a thermodynamic context by Zurek [13], where he related it to the advantage a quantum Maxwell demon could have over a classical one. In general the quantum discord,  $\delta(A : S)$ , is defined as the minimum of  $\delta^{\mathcal{P}}(A : S)$  over all sets of projectors  $\{\Pi_k^{\mathcal{P}}\}_k$  whereas in our case this set is fixed (see e.g. Modi *et al.* [52] for a review). Therefore it is found that even for states with no quantum discord, usually referred to as classically correlated states, a difference in work associated with thermodynamic projection processes can be observed. This contrasts with the erasure process [14] where an advantage could only be gained for highly entangled states.

One may ask what global states on  $SA$  maximize  $\langle W_{\text{opt}} \rangle$  for a given state  $\rho^S$  on  $S$ . Expectedly, it can be shown that purifications of  $\rho^S$  yield the best improvement in terms of extracted work. Given  $\rho^S = \sum_l a_l |l\rangle\langle l|$  any purification is, up to isometries on the purifying system [17], equivalent to  $|\Psi\rangle = \sum_l \sqrt{a_l} |l\rangle^S |l\rangle^A$  for some orthonormal basis  $\{|l\rangle^A\}_l$  of  $A$ . For such a state the conditional states on  $A$  after the projection,  $\rho_k^A$ , are pure for all  $k$  which implies that they have zero entropy. This implies

$$\begin{aligned}\delta^{\mathcal{P}}(A : S) &= S(\rho^S) - S(|\Psi\rangle\langle\Psi|) + \sum_k p_k S(\eta_k^A) \\ &= S(\rho^S) - 0 + \sum_k p_k \cdot 0 = S(\rho^S).\end{aligned} \quad (\text{H5})$$

The optimal total extracted work from a purified state on  $SA$  in a thermodynamic projection process is therefore  $\langle W_{\text{opt}} \rangle = k_B T S(\eta^S) - \Delta U^{\mathcal{P}}$  which can be shown to be

the maximum for fixed  $\rho^S$  and projectors  $\{\Pi_k^{\mathcal{P}}\}_k$ . One way to see this is the following Lemma [53].

**Lemma 1.** *Let  $\rho^{SA}$  be an arbitrary state on a bipartite system  $SA$ , and let  $\{\Pi_k^{\mathcal{P}}\}_k$  be a complete set of rank-1 orthogonal projectors on  $S$ . Consider the global state after the projection,  $\eta^{SA} = \sum_k (\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A) \rho^{SA} (\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A) = \sum_k p_k \Pi_k^{\mathcal{P}} \otimes \eta_k^A$ , where  $p_k = \text{tr}[(\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A) \rho^{SA}]$  are the probabilities to measure  $k$  on  $S$  and  $\eta_k^A = p_k^{-1} \text{tr}_S[(\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A) \rho^{SA} (\Pi_k^{\mathcal{P}} \otimes \mathbb{1}^A)]$  are the conditional states on  $A$ . Then*

$$S(\rho^{SA}) \geq \sum_k p_k S(\eta_k^A). \quad (\text{H6})$$

*Proof.* We model the process on  $S$  as an isometry  $\Phi^{S \rightarrow S\tilde{S}} = \sum_k |\psi_k\rangle^{\tilde{S}} \otimes \Pi_k^S$ , where  $\tilde{S}$  is a copy of  $S$  and  $\{|\psi_k\rangle^{\tilde{S}}\}_k$  is an orthonormal basis of  $\tilde{S}$ . The state after applying the isometry is denoted  $\eta^{S\tilde{S}A} = \Phi \rho^{SA} \Phi^\dagger$  and we note that  $\text{tr}_{\tilde{S}}[\eta^{S\tilde{S}A}] = \eta^{SA}$ . Furthermore, isometries do not change (von Neumann) entropy and thus,  $S(\eta^{S\tilde{S}A}) = S(\rho^{SA})$ . In addition, by construction of  $\tilde{S}$  the marginals of the final state on  $S$  and  $\tilde{S}$  have the same entropy:  $S(\eta^{\tilde{S}}) = S(\eta^S)$ . Since  $S(\eta^{S\tilde{S}A}) \geq |S(\eta^{SA}) - S(\eta^{\tilde{S}})| \geq S(\eta^{SA}) - S(\eta^{\tilde{S}})$  (see e.g. Nielsen & Chuang [17]). Thus

$$\begin{aligned}S(\rho^{SA}) &= S(\eta^{S\tilde{S}A}) \geq S(\eta^{SA}) - S(\eta^{\tilde{S}}) \\ &= S(\eta^{SA}) - S(\eta^S) = \sum_k p_k S(\eta_k^A),\end{aligned} \quad (\text{H7})$$

where in the last equality we made use of the fact that  $\eta^{SA}$  is a classical-quantum state.  $\square$

Going back to Eq. (H3) and applying the the above Lemma we see that in general  $\langle W_{\text{opt}} \rangle = k_B T \Delta \mathcal{S}^{\mathcal{P}} - \Delta U^{\mathcal{P}} = k_B T [S(\eta^S) + \sum_k p_k S(\eta_k^A) - S(\rho^{SA})] - \Delta U^{\mathcal{P}} \leq k_B T S(\eta^S) - \Delta U^{\mathcal{P}}$ , which proves that purifications on  $SA$  yield the maximally possible extracted work.

- 
- [1] Landauer, R. Dissipation and heat generation in the computing process. *IBM J. Res. Develop.* **5**, 148-156 (1961).
  - [2] Scully, M. O., Zubairy, M.S., Agarwal, G.S. & Walther, H. Extracting Work from a Single Heat Bath via Vanishing Quantum Coherence. *Science* **299**, 862 (2003).
  - [3] Kosloff, R. & Levy, A. Quantum Heat Engines and Refrigerators: Continuous Devices. *Annu. Rev. Phys. Chem.* **65**, 365 (2014).
  - [4] Roßnagel, J., Abah, O., Schmidt-Kaler, F., Singer, K. & Lutz, E. Nanoscale Heat Engine Beyond the Carnot Limit. *Phys. Rev. Lett.* **112**, 030602 (2014).
  - [5] Mukamel, S. Quantum extension of the Jarzynski relation: Analogy with stochastic dephasing. *Phys. Rev. Lett.* **90**, 170604 (2003).
  - [6] Talkner, P., Lutz, E. & P. Hänggi. Fluctuation theorems: Work is not an observable. *Phys. Rev. E* **75**, 050102 (R) (2007).
  - [7] Janzing, D., Wocjan, P., Zeier, R., Geiss, R. & Beth, T. Thermodynamic cost of reliability and low temperatures: tightening Landauer's principle and the second law. *Int. J. Theor. Phys.* **39**, 2717 (2000).
  - [8] Brandão, F.G.S.L., Horodecki, M., Ng, N.H.Y., Oppenheim, J. & Wehner, S. The second laws of quantum thermodynamics. *PNAS* **112**, 3275 (2015).
  - [9] Lostaglio, M., Jennings, D. & Rudolph, T. Description of quantum coherence in thermodynamic processes requires constraints beyond free energy. *Nat. Commun.* **6**, 6383 (2015).



- [10] Jacobs, K. Quantum measurement and the first law of thermodynamics: The energy cost of measurement is the work value of the acquired information. *Phys. Rev. E* **86** 040106(R) (2012).
- [11] Erez, N. Thermodynamics of projective quantum measurements. *Phys. Scr.* **151**, 014028 (2012).
- [12] Åberg, J. Catalytic Coherence. *Phys. Rev. Lett.* **113**, 1504022 (2014).
- [13] Zurek, W.H. Quantum discord and Maxwell's demons. *Phys. Rev. A* **67**, 012320 (2003).
- [14] del Rio, L., Åberg, J., Renner, R., Dahlsten, O. & Vedral, V. The thermodynamic meaning of negative entropy. *Nature* **474**, 61 (2011).
- [15] Erez, N., Gordon, G., Nest, M. & Kurizki, G. Thermodynamic control by frequent quantum measurements. *Nature* **452**, 724 (2008).
- [16] Zurek, W. H. Decoherence, einselection, and the quantum origins of the classical. *Rev. Mod. Phys.* **75**, 715 (2003).
- [17] Nielsen, M.A. & Chuang, I.L. *Quantum Computation and Quantum Information* Cambridge University Press, Cambridge UK (2000).
- [18] Streater, R.F. Convergence of the Quantum Boltzmann Map. *Commun. Math. Phys.* **98**, 177 (1985).
- [19] Liphardt, J., Dumont, S., Smith, S.B., Tinoco, I. & Bustamante, C. Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality. *Science* **296**, 1832 (2002).
- [20] M. Campisi, P. Talkner, and P. Hänggi. Fluctuation theorem for arbitrary open quantum systems. *Phys. Rev. Lett.* **102**, 210401 (2009).
- [21] Bérut, A., Arakelyan, A., Petrosyan, A., Ciliberto, S., Dillenschneider, R. & Lutz, E. Experimental verification of Landauer's principle linking information and thermodynamics. *Nature* **483**, 7388 (2012).
- [22] Saira, O-P., *et al.* Test of the Jarzynski and Crooks Fluctuation Relations in an Electronic System. *Phys. Rev. Lett.* **109**, 180601 (2012).
- [23] Skrzypczyk, P., Short, A.J. & Popescu, S. Work extraction and thermodynamics for individual quantum systems. *Nat. Commun.* **4**, 4185 (2013).
- [24] Esposito, M., Lindenberg, K. & Van den Broeck, Ch. Entropy production as correlation between system and reservoir. *New J. Phys.* **12**, 013013 (2010).
- [25] Anders, J. & Giovanetti, V. Thermodynamics of discrete quantum processes. *New J. Phys.* **15**, 033022 (2013).
- [26] Batalhao, T.B. *et al.* Experimental Reconstruction of Work Distribution and Study of Fluctuation Relations in a Closed Quantum System. *Phys. Rev. Lett.* **113**, 140601 (2014).
- [27] Morley, G.W., *et al.* The initialization and manipulation of quantum information stored in silicon by bismuth dopants. *Nat. Mat.* **9**, 725 (2010).
- [28] Maruyama, K., Nori, F. & Vedral, V. Colloquium: The physics of Maxwell's demon and information. *Rev. Mod. Phys.* **81**, 1 (2009).
- [29] Åberg, J. Truly work-like work extraction via a single-shot analysis. *Nat. Commun.* **4**, 1925 (2013).
- [30] Horodecki, M. & Oppenheim, J. Fundamental limitations for quantum and nanoscale thermodynamics. *Nat. Commun.* **4**, 2059 (2013).
- [31] Korzekwa, K., Lostaglio, M., Oppenheim, J. & Jennings, D. The extraction of work from quantum coherence. arXiv:1506.07875 (2015).
- [32] Henderson, L. & Vedral, V. Classical, quantum and total correlations. *J. Phys. A: Math. Gen.* **34**, 6899 (2001).
- [33] Maunz, P. Gentle measurement. *Nature* **475**, 180 (2011);
- [34] Volz, J., Gehr, R., Dubois G., Estève, J. & Reichel, J. Measurement of the internal state of a single atom without energy exchange. *Nature* **475**, 210 (2011).
- [35] Thanks to Michael Wolf who pointed this out to us.
- [36] Appleby, D.M. *J. Math. Phys.* **46**, 052107 (2005).
- [37] Durt, T., Englert, B.-G., Bengtsson, I. & Życzkowski, K. *Int. J. Quantum Inform.* **08**, 535 (2010).
- [38] Ando, T. *Linear Algebra Appl.* **118**, 163-248 (1989).
- [39] Gemmer, J., Michel, M. & Mahler, G. *Quantum Thermodynamics* Springer (2009).
- [40] Pusz, W. & Woronowicz, S.L. Passive states and KMS states for general quantum systems. *Comm. Math. Phys.* **58**, 273 (1978).
- [41] Lenard, A. Thermodynamical proof of the Gibbs formula for elementary quantum systems. *J. Stat. Phys.* **19**, 575 (1978).
- [42] Morikuni, Y. & Tasaki, H. Quantum Jarzynski-Sagawa-Ueda Relations. *J. Stat. Phys.* **143**, 1 (2011).
- [43] Seifert, U. Stochastic thermodynamics, fluctuation theorems and molecular machines. *Rep. Prog. Phys.* **75**, 126001 (2012).
- [44] Brandão, F.G.S.L., Horodecki, M., Oppenheim, J., Renes, J.M. & Spekkens, R.W. *Phys. Rev. Lett.* **111**, 250404 (2013).
- [45] Datta, N. Min- and max-relative entropies and a new entanglement monotone. *IEEE Trans. Inf. Theory* **55**, 2816-2826 (2009).
- [46] Tomamichel, M. *A framework for non-asymptotic quantum information theory*. PhD Thesis, ETH Zurich (2012).
- [47] Dupuis, F., Kraemer, L., Faist, P., Renes, J.M. & Renner, R. XVIIth International Congress on Mathematical Physics, 134-153 (2013).
- [48] Mazzola, L., De Chiara, G. & Paternostro, M. Measuring the characteristic function of the work distribution. *Phys. Rev. Lett.* **110** 230602 (2013).
- [49] Roncaglia, A., Cerisola, F. & Paz, J.P. *Phys. Rev. Lett.* (2014).
- [50] Shannon, C.E. *Bell Syst. Tech. J.* **27**, (3) 379-423 (1948)
- [51] Ollivier, H. & Zurek, W.H. Quantum discord: a measure of the quantumness of correlations. *Phys. Rev. Lett.* **88**, 017901 (2000).
- [52] Modi, K., Brodutch, A., Cable, H., Paterek, T. & Vedral, V. *Rev. Mod. Phys.* **84**, 1655 (2012).
- [53] Renner, R. Private Communication (2014).